AD NUMBER

ADB262564

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies only; Proprietary Info.; Oct 99. Other requests shall be referred to U.S. Army Medical Research and Materiel Command, 504 Scott St., Fort Detrick, MD 21702-5012.

AUTHORITY

USAMRMC ltr, 1 Apr 2003

AD_____

Award Number: DAMD17-98-1-8486

TITLE: Unbiased Outcome Estimates for Conservative vs. Aggressive Treatment of Early Stage Prostate Cancer from Retrospective Data: An Instrumental Variables Approach

PRINCIPAL INVESTIGATOR: Elizabeth A. Chrischilles, Ph.D., M.S.

CONTRACTING ORGANIZATION: University of Iowa
Iowa City, Iowa 52242-1320

REPORT DATE: October 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

NOTICE

USING GOVERNMENT DRAWINGS, SPECIFICATIONS, OR OTHER DATA INCLUDED IN THIS DOCUMENT FOR ANY PURPOSE OTHER THAN GOVERNMENT PROCUREMENT DOES NOT IN ANY WAY OBLIGATE THE U.S. GOVERNMENT. THE FACT THAT THE GOVERNMENT FORMULATED OR SUPPLIED THE DRAWINGS, SPECIFICATIONS, OR OTHER DATA DOES NOT LICENSE THE HOLDER OR ANY OTHER PERSON OR CORPORATION; OR CONVEY ANY RIGHTS OR PERMISSION TO MANUFACTURE, USE, OR SELL ANY PATENTED INVENTION THAT MAY RELATE TO THEM.

LIMITED RIGHTS LEGEND

Award Number: DAMD17-98-1-8486
Organization: University of Iowa
Location of Limited Rights Data (Pages):

Those portions of the technical data contained in this report marked as limited rights data shall not, without the written permission of the above contractor, be (a) released or disclosed outside the government, (b) used by the Government for manufacture or, in the case of computer software documentation, for preparing the same or similar computer software, or (c) used by a party other than the Government, except that the Government may release or disclose technical data to persons outside the Government, or permit the use of technical data by such persons, if (i) such release, disclosure, or use is necessary for emergency repair or overhaul or (ii) is a release or disclosure of technical data (other than detailed manufacturing or process data) to, or use of such data by, a foreign government that is in the interest of the Government and is required for evaluational or informational purposes, provided in either case that such release, disclosure or use is made subject to a prohibition that the person to whom the data is released or disclosed may not further use, release or disclose such data, and the contractor or subcontractor or subcontractor asserting the restriction is notified of such release, disclosure or use. This legend, together with the indications of the portions of this data which are subject to such limitations, shall be included on any reproduction hereof which includes any part of the portions subject to such limitations.

THIS TECHNICAL REPORT HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION.

Narasingha Chareen Munroe
12/22/00

_____

_____

_____

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1999	Annual (1-Oct-98 - 30-Sep-99)

4. TITLE AND SUBTITLE

Unbiased Outcome Estimates for Conservative vs. Aggressive Treatment of Early Stage Prostate Cancer from Retrospective Data: An Instrumental Variables Approach

5. FUNDING NUMBERS

DAMD17-98-1-8486

6. AUTHOR(S)

Elizabeth A. Chrischilles, Ph.D., M.S.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Iowa
Iowa City, Iowa 52242-1320

E-MAIL:
e-chrischilles@uiowa.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution authorized to U.S. Government agencies only (proprietary information, Oct 99). Other requests for this document shall be referred to U.S. Army Medical Research and Materiel Command, 504 Scott Street, Fort Detrick, Maryland 21702-5012.

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

The purpose of this study is to use instrumental variables techniques to estimate the outcome differences between aggressive treatment and conservative management among marginal patients with early stage prostate cancer; estimate true cost-effectiveness ratios; and determine whether and what type of patients may be safely shifted from aggressive to conservative treatment.

In this annual report we describe the factors that are related to the sorting of patients to receive conservative or aggressive treatments. We have demonstrated that selected access to care variables (the candidate instrumental variables) group patients such that the groups have different rates of aggressive treatment but do not differ meaningully in demographic, tumor, or comorbidity characteristics. Thus, these variables appear to satisfy the two criteria that an instrument must fulfill: (1) the variable must be related to the possibility of patients receiving a particular treatment; and (2) the variable must have no effect on outcomes either directly or indirectly (e.g., through relationships with unmeasured confounding factors such as patient severity and unrecorded treatments). If the candidate instruments had failed this test, it would not have been possible to proceed to Task 2, estimating unbiased treatment effects for conservative vs. aggressive treatment.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Prostate Cancer	37
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Limited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

___ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

___ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

N/A In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

N/A In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

N/A In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

N/A In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Elizabeth A. Chrachko 10-29-99

PI - Signature Date

Table of Contents

Introduction

Prostate cancer is the most frequent cancer among American men and is the second leading cause of cancer-related deaths in all males.[1] With the advent of widespread screening with prostate-specific antigen (PSA), increasing numbers of men have been diagnosed with asymptomatic, localized, prostate cancer.[2] Among patients with clinically localized disease it is not known whether conservative management, i.e., "watchful waiting" or aggressive treatment, i.e., radiation therapy or radical prostatectomy, has better effectiveness. This is because men who are diagnosed with early stage prostate cancer may die of other causes before prostate cancer progresses enough to affect health. Both radical prostatectomy and radiation treatment have high rates of complications such as sexual impotence, urinary incontinence, and infection which adversely affect health. There is also a risk of surgical mortality with prostatectomy. Ideally, clinicians would identify men whose life expectancy was short enough that their prostate cancer would not be expected to progress substantially in their remaining lifetime. These men would receive conservative treatment (and no complications from aggressive treatment). For the rest, the benefit of aggressive treatment would be worth the risk of complications and they would receive aggressive treatment. However, although current prognostic factors for prostate carcinoma provide important information for patient care, the ideal method with which to incorporate the information attained from tumor-related factors (clinical stage, histologic grade, and PSA level), patient age, and comorbidity into a manageable prognostic score has not been found. The purpose of this study is to use instrumental variables techniques to estimate the outcome differences between aggressive treatment and conservative management among marginal patients with clinically localized disease; combine the health outcome and cost estimates to estimate true cost-effectiveness ratios; and using measured characteristics such as patient age, tumor grade, and the extent of co-morbid conditions, determine whether and what type of patients may be safely shifted from aggressive to conservative treatment.

Body

Five tasks were outlined in the Statement of Work for this project. Of these, only the first task was the subject of the first project year (October 1, 1998 through September 30, 1999). Tasks 2 through 5 begin in months 13 or later of the project. This annual report describes results of task #1 from the approved Statement of Work. The focus of task #1 is on describing the factors associated with treatment choice in early stage prostate cancer. The associated work product will be a published manuscript. The exact language of task #1 follows:

Task 1. Describe the factors that are related to the sorting of patients into conservative or aggressive treatments, Months 1-15.

a. Obtain data from SEER-HCFA linked databases and AMA Master File (Months 1-2).
b. Create analytic files (Months 3-4).
c. Construct and validate instrumental variables (Months 5-6).
d. Construct and validate treatment variables (Months 5-6).
e. Conduct analysis (Months 7-12) Examine patient-specific factors (demographic, co-morbidity, and tumor-related) and a series of factors related to treatment variation and theoretically unrelated to unmeasured confounders (candidate instrumental variables).
f. Prepare and submit manuscript (Months 13-15).

1.0 Databases Acquired

The primary data required for this study have been acquired and include:

(1) Medicare data files merged with SEER Program data (the SEER-Medicare linked data) for *all* SEER Program sites;

(2) A list of all radiation treatment centers providing service in the region containing each registry, including zip code of location and years in operation; and

(3) Area Resource File (ARF) of area provider counts.

The AMA Master File is no longer needed because the ARF will provide the area provider counts. With respect to the SEER-Medicare linked data (number (1) above), two particular data files were used for the analyses in this report. These were the Patient Entitlement and Diagnosis Summary File (PEDSF) and the Medicare Provider Analysis and Review files (MEDPAR). The SEER data used in SEER-Medicare projects are contained in a customized file known as the Patient Entitlement and Diagnosis Summary File (PEDSF). This file contains one record per person for individuals in the SEER Program database who have been matched with Medicare enrollment records. For persons appearing in the PEDSF file, basic SEER Program diagnostic information is available for up to 10 diagnosed cancer cases. Data also include Medicare entitlement data for the person for the period 1984-1994. The MEDPAR file includes Medicare data about all short stays, long stays and skilled nursing facility (SNF) records for each calendar year. There is one summarized record per hospitalization, including up to 10 ICD-9 diagnoses and 10 ICD-9 procedures provided during the hospitalization. Bills received from SNFS are included even if there is no discharge date as for many persons in SNFs, there is no reported discharge date. The MEDPAR file needs to be subset if only short stay hospital records are needed.

We obtained the SEER-Medicare linked data for Iowa before the start of the study so were able to begin preliminary analyses on October 1, 1998. On October 19, 1998 we requested the Medicare data files merged with SEER Program data (the SEER-Medicare linked data) for *all* SEER Program sites from the Applied Research Branch of the National Cancer Institute. We received these data in February, 1999. We also obtained from each SEER registry a list of all radiation treatment centers providing service in the area covered by the registry. The lists contained the zip code of each center and years between 1984 and 1995 that each center provided services. We have also obtained the Area Resource File (ARF).

2.0 Data Validity Studies

Data quality was examined in a series of analyses. First, we programmed a case selection process for use with the SEER-Medicare linked data and evaluated it using the full Iowa SEER database. Then, for the prostate cancer cases selected from the 11 SEER registries via the SEER-Medicare linked database we evaluated the quality of the death date variables and the quality of the AJCC cancer staging variable. The results of these data validity studies are reported here.

2.1 Validation of the Case Selection Algorithm

Initially, we used the Iowa SEER-Medicare linked data in order to develop and validate our case selection rules and to conduct preliminary data quality studies of key variables. The SEER-Medicare

linked data include a restricted set of all the available SEER data. Because we additionally had access to the full Iowa SEER database we were able to treat the full database as the "gold standard" against which to compare the cases of early stage prostate cancer selected using only the SEER-Medicare linked data.

Using the Iowa SEER-Medicare database and also using the full Iowa SEER database, two database analysts independently applied the study inclusion and exclusion criteria. The cases selected from each source were then compared. The inclusion and exclusion criteria are listed below:

Included in the study will be:

1. Men
2. Resident of incidence reporting area (place of residence code 1) at time of diagnosis
3. First primary cancer only (sequence number 00 or 01 or larger sequence number if prior sequence numbers are in situ)
4. age 65 and older at diagnosis
5. malignant prostate cancer of any histologic type (except small cell and soft tissue sarcomas)
6. diagnosed between January 1, 1985 and December 31, 1993
7. because more advanced disease may not be recognized until the time of prostatectomy, we will initially include cancers of any stage.

Excluded from study will be:

1. men under age 65 at the time of diagnosis
2. non-resident of incidence reporting area at time of diagnosis
3. cases diagnosed before 1985
4. in situ prostate cancer
5. men who have had a prior cancer diagnosis (sequence code > 01)
6. small cell and soft tissue sarcomas will be included in the initial data run but may be excluded after counting the frequency of occurrence.

In the Iowa validation study there were 12,532 patients selected from the PEDSF file according to the above criteria (Tables 1-3). Also based on the same criteria there were 13,161 patients selected from the full Iowa SEER database (Table 4).

Table 1. Number of Patients Selected from the SEER-Medicare Linkage Data, Patient Entitlement and Diagnosis Summary File (PEDSF), According to Diagnosis Context of the Prostate Cancer Diagnosis and Case Number Position of the Prostate Cancer Diagnosis.

Diagnosis Context	**Case #1**	**Case #2**	**Case #3**	**Total(%)**
1st Primary	12,140	0	0	12,140(96.87)
In Situ Prostate Before Malignant Prostate	0	0	0	0
Other In Situ Before Malignant Prostate	0	65	1	66(0.53)
Simultaneous Malignant with Malignant Prostate	150	156	7	313(2.5)
Simultaneous In Situ with Malignant prostate	7	3	0	10(0.08)
Combination of the above	0	0	3	3(0.02)
Total(%)	12,297(98.13)	224(1.78)	11(0.09)	12,532(100)

Table 2. Number of Patients Selected from Patient Entitlement and Diagnosis Summary File (PEDSF), By Histologic Stage.

Stage:	Unstaged	Distant	Localized	Regional	Total
Number of Patients (%):	1,365(10.89)	1,787(14.26)	7,145(57.01)	2,235(17.83)	12,532(100)

Table 3. Number of Patients Selected from Patient Entitlement and Diagnosis Summary File (PEDSF), By Diagnosis Context of Prostate Cancer and Whether Cancer Record Was Also Selected From the Full Iowa SEER Database.

Diagnosis Context	Selected in Iowa SEER Database	Not selected in Iowa SEER Database	Total(%)
1st Primary	12,078	62	12,140(96.87)
In Situ Prostate Before Malignant Prostate	0	0	0
Other In Situ Before Malignant Prostate	60	6	66(0.53)
Simultaneous Malignant with Malignant Prostate	303	10	313(2.50)
Simultaneous In Situ with Malignant prostate	10	0	10(0.08)
Combination of the above	2	1	3(0.02)
Total	12,453(99.37)	79(0.63)	12,532(100)

Table 4. Number of Patients Selected from the Full Iowa SEER Database, By Diagnosis Context and Whether Cancer Record Was Also Selected From the Patient Entitlement and Diagnosis Summary File (PEDSF).

	Found in PEDSF		Not Found in PEDSF	Total
	Selected	Not Selected		
1st Primary	12,068	3	684	12,755(96.92)
In Situ Prostate Before Malignant Prostate	0	0	0	0
Other In Situ Before Malignant Prostate	70	4	3	77(0.6)
Simultaneous Malignant with Malignant Prostate	302	2	12	316(2.4)
Simultaneous In Situ with Malignant prostate	11	0	0	11(0.08)
Combination of the above	2	0	0	2(0.02)
Total	12,453(94.62)	9(0.08)	699(5.3%)	13,161(100)

Of the 12,532 patients selected from the PEDSF, 12,453 (99.37%) were also selected from the full Iowa SEER database. The 79 patients (0.63%) who were not selected from the full Iowa SEER database were actually found in the database but were excluded according to the same criteria (Tables 5 and 6). One possible reason is that the full Iowa SEER database has more up-to-date information than the PEDSF file.

Table 5. Number of Patients Selected from the Patient Entitlement and Diagnosis Summary File (PEDSF) But Not Selected From the Iowa SEER Database, By Diagnosis Context and Histologic Stage (Obtained From the PEDSF).

	Histologic Stage				
Diagnosis Context	**Localized**	**Regional**	**Distant**	**Unstaged**	**Total(%)**
1st Primary	32	11	8	11	62(78.48)
Other Insitu Before Malignant Prostate	3	1	1	1	6(7.59))
Simultaneous Malignant with Malignant Prostate	10	0	0	0	10(12.66)
Some Combination of above	1	0	0	0	1(1.27)
Total(%)	46 (58.23)	12(15.19)	9(11.39)	12(15.19)	79(100)

Table 6. Number of Patients Selected from the Patient Entitlement and Diagnosis Summary File (PEDSF) But Not Selected From the Iowa SEER Database, By Age Category and Year of Diagnosis (Obtained From the PEDSF).

	Year of Diagnosis									
Age	**1985**	**1986**	**1987**	**1988**	**1989**	**1990**	**1991**	**1992**	**1993**	**Total**
65~69	2	5	5	6	3	1	4	11	6	43 (54.43)
70~74	0	0	1	1	0	3	2	1	3	11 (13.92)
75~79	1	0	2	1	2	0	2	2	1	11 (13.92)
80~84	0	1	1	2	1	0	2	1	0	8 (10.13)
85~89	2	0	1	1	1	0	1	0	0	6 (7.59)
Total (%)	5 (6.33)	6 (7.59)	10 (12.66)	11 (13.92)	7 (8.86)	4 (5.06)	11 (13.92)	15 (18.99)	10 (12.66)	79 (100)

Of the 13,161 patients selected from the full Iowa SEER database, 94.62% (12,453 patients) were also selected from the PEDSF. The difference, 708 patients, was accounted for by 0.08% (9 patients) who were found in the PEDSF file but were not selected according to the same criteria (Table 4) and 5.31% (699 patients) who were not found in the PEDSF file at all (Tables 7 and 8).

Table 7. Number of Patients Selected from the Iowa SEER Database but Not Selected from the Patient Entitlement and Diagnosis Summary File, By Diagnosis Context and Histologic Stage (Obtained from the Iowa SEER Database).

	Histologic Stage				
	Localized	**Regional**	**Distant**	**Unstaged**	**Total(%)**
1st Primary	85	293	246	60	684(97.85)
Other Insitu Before Malignant Prostate	1	2	0	0	3(0.43)
Simultaneous Malignant with Malignant Prostate	2	4	6	0	12(1.72)
Total(%)	88(12.59)	299(42.78)	252(36.05)	60(8.58)	699(100)

Table 8. Number of Patients Selected from the Iowa SEER Database but Not Selected from the Patient Entitlement and Diagnosis Summary File, By Age Category and Year of Diagnosis (Obtained from the Iowa SEER Database).

Age	Year of Diagnosis									
	1985	1986	1987	1988	1989	1990	1991	1992	1993	Total
65~69	9	6	7	7	17	5	6	17	32	106 (15.16)
70~74	3	7	8	14	21	6	12	15	39	125 (17.88)
75~79	14	4	14	16	29	6	15	25	55	178 (25.46)
80~84	3	7	15	11	22	6	11	17	44	136 (19.46)
85~89	5	4	5	7	13	9	10	13	29	95 (13.59)
90~94	0	0	8	6	8	2	4	4	12	44 (6.29)
95~99	0	0	0	0	4	0	1	3	6	14 (2.00)
100 & OLDER	0	0	0	0	1	0	0	0	0	1 (0.14)
Total (%)	34 (4.86)	28 (4.01)	57 (8.15)	61 (8.73)	115 (16.45)	34 (4.86)	59 (8.44)	94 (13.45)	217 (31.04)	699 (100)

In summary, the positive predictive value of the SEER-Medicare selection process was 99.37% and the sensitivity of the process was 94.62%. If we assume that the 699 SEER-detected prostate cancers that were not found in the SEER-Medicare database were not Medicare eligible or non-linked, the sensitivity of the selection process for detecting early stage, first primary, prostate cancer among Medicare eligible and linked men, is nearly 100%.

As a final validation step we examined the consistency of tumor grade and diagnosis context for the 12,453 matched cases. There were 12,433 patients (2781+5496+2882+325+1+948=12,433), or 99.84% who had equivalent tumor grade in both databases (Table 9). Also, 99.8% (12063+60+298+8+2=12,431) had equivalent diagnosis context in both databases (Table 10).

Table 9. Number of Matched Patients By Grade in the Patient Entitlement and Diagnosis Summary File and in the Iowa SEER Database.

Tumor Grade in PEDSF	Tumor Grade in Iowa SEER Database						
	I: Well differen-tiated	II: Moderately differen-tiated	III: Poorly Differen-tiated	IV: Un-differen-tiated	B-cell; Pre-B	Grade Unknown	Total (%)
I: Well Differentiated	2,781	6	0	0	0	1	2,788 (22.39)
II: Moderately Differentiated	1	5,496	3	0	0	1	5,501 (44.17)
III: Poorly Differentiated	0	5	2,882	1	0	0	2,888 (23.19)
IV:Un-differentiated	0	0	0	325	0	0	325 (2.61)
B-cell;Pre-B; B-precursor	0	0	0	0	1	0	1 (0.01)
Grade unknown	0	1	1	0	0	948	950 (7.63)
TOTAL (%)	2,782 (22.34)	5,508 (44.23)	2,886 (23.18)	326 (2.62)	1 (0.01)	950 (7.63)	12,453 (100)

Table 10. Number of Matched Patients By Diagnosis Context in the Patient Entitlement and Diagnosis Summary File and in the Iowa SEER Database.

Diagnosis Context in PEDSF	Diagnosis Context in Iowa SEER Database						
	1st Primary	Insitu Prostate Before Malignant Prostate	Other Insitu Before Malignant Prostate	Simultaneous Malignant with Malignant Prostate	Simultaneous Insitu with Malignant prostate	Combination of the left columns	Total (%)
1st Primary	12,063		10	2	3		12,078 (96.99)
Insitu Prostate Before Malignant Prostate							0
Other Insitu Before Malignant Prostate			60				60 (0.48)
Simultaneous Malignant with Malignant Prostate	5			298			303 (2.43)
Simultaneous Insitu with Malignant prostate				2	8		10 (0.08)
Combination of the above						2	2 (0.02)
Total (%)	12,068 (96.91)	0	70 (0.56)	302 (2.43)	11 (0.09)	2 (0.02)	12,453 (100)

2.2 Mortality Validity Check

Medicare data and SEER data each contain a death date, obtained via different mechanisms. We compared these two fields and found good concordance. Of the 97,896 subjects meeting exclusion and inclusion criteria from all 11 SEER registries, 64.38% (63,020 patients) had neither Medicare death dates nor SEER death dates while 35.62% (34,876 patients) had death dates in Medicare or SEER data (Table 11). Among the 34,876 patients who had death dates in Medicare or SEER, 3.4% had death dates in Medicare records but not in SEER, 2.43% had death dates in SEER but not in Medicare, and the remaining 94.17% (32,841 patients) had death dates in both data sets. Of those 32,841 patients who had death dates in both data, 99.7% (32,744 patients) had the same years of death recorded in both. (However, a few of them have different months.) Looking by year of death, 92% of Medicare-only death dates fell in 1995 and 71% of the SEER only death dates fell in 1993/1994 (Table 12).

Table 11.Death Dates of Eligible Subjects in the Medicare and SEER Data.

Death in Medicare only				**1,187 (3.40%)**
Death in SEER only				**848 (2.43%)**
Death in both Medicare and SEER	Same year of death	Same month	32,501 (99.26%)	**32,744 (93.89%)**
		Different month	243 (0.74%)	
		Subtotal	32,744 (100%)	
	Different year of death	1 year	79 (81.4%)	**97 (0.28%)**
		2 year	9 (9.3%)	
		3 year	5 (3.2%)	
		4 year	2 (2.06%)	
		5 year	2 (2.06%)	
		Subtotal	97 (100%)	
Total				**34,786 (100%)**

Table 12.Distributions of Subjects Who Have Inconsistent Death Data in the Two Data Sources, By Year of Death Recorded in the Data.

Year of Death	Medicare Death Only	SEER Death Only
1985		5(0.59%)
1986		11(1.30%)
1987	4(0.34%)	17(2%)
1988	4(0.34%)	33(3.89%)
1989	4(0.34%)	33(3.89%)
1990	5(0.42%)	40(4.72%)
1991	12(1.01%)	63(7.43%)
1992	11(0.93%)	38(4.48%)
1993	25(2.11%)	253(29.83%)
1994	30(2.53%)	355(41.86%)
1995	1092(92%)	
Total	**1,187(100%)**	**848(100%)**

2.3 Quality of AJCC Stage Variable

The PEDSF file has information for the AJCC stage of all cancer records in all the covered years (1985-1993). For the years before 1988, the stages were derived from the 4 digit SEER extent of

disease (EOD) codes. For the years after 1988, the stages were derived from the 10 digit EOD codes. Tables 13-15 examine the AJCC variable. Overall, 38.05% (37,247 patients) of the selected patients had unstaged tumors by the AJCC system (Table 13). In Table 14, it can be seen that all those AJCC unstaged patients either had histologically localized cancers (23,803 patients) or histologically unstaged cancers (13,444 patients). Of the AJCC unstaged tumors, those that were also histologically unstaged were diagnosed from 1985 through 1993, while the 23,803 localized (histogically) but AJCC unstaged patients were all diagnosed after 1987 (Table 15). The same 23,803 patients constitute 41.5% of all the histogically localized cancer patients (57,325) (Table 14). The percent of prostate tumors that are unstaged is known to have risen in the era of PSA testing, to about 15%. However, the abrupt increase in AJCC unstaged percent in 1988 (and the corresponding abrupt decrease in stage I and II cancers) is far in excess of this amount. The algorithm to derive AJCC stage from the SEER EOD codes can be applied very stringently or there is a more liberal way of applying the algorithm. It appears that the stringent version has been used, resulting in "unstaged" being assigned when the exact details required by the algorithm are not available. The AJCC stage variable will not be usable for this project. Instead we will rely on the SEER histologic stage.

Table 13. Number of Prostate Canceer Patients By AJCC Stage and Year of Diagnosis.

	Year of Diagnosis (column %)									
AJCC Stage	**1985**	**1986**	**1987**	**1988**	**1989**	**1990**	**1991**	**1992**	**1993**	**Total**
Stage 0	.	.	.	574 (7.25)	589 (7.12)	695 (6.48)	715 (5.20)	744 (3.71)	633 (3.76)	3,950 (4.03)
Stage I	1058 (16.9)	1228 (18.57)	1528 (20.35)	942 (11.90)	959 (11.59)	1280 (11.94)	1778 (12.94)	2483 (12.40)	2503 (14.87)	13,759(14.05) Stage I
Stage II	2536 (40.51)	2545 (38.48)	2982 (39.71)	588 (7.43)	721 (8.71)	1042 (9.72)	1475 (10.73)	2156 (10.76)	1768 (10.51)	15,813(6.15) Stage II
Stage III	517 (8.26)	556 (8.41)	660 (8.79)	672 (8.49)	904 (10.92)	1224 (11.42)	1805 (13.13)	2867 (14.31)	2464 (14.64)	11,669(11.92) **Stage III**
Stage IV	1503 (24.01)	1544 (23.34)	1540 (20.51)	1591 (20.10)	1626 (19.65)	1799 (16.78)	1917 (13.95)	2240 (11.18)	1698 (10.09)	15,458(15.79) **Stage IV**
UNSTAGED	646 (10.32)	741 (11.20)	799 (10.64)	3550 (44.84)	3476 (42.01)	4680 (43.66)	6055 (44.05)	9539 (47.63)	7761 (46.12)	37,247(38.05) Unstaged
Total (row %)	**6,260** **(6.39)**	**6,614** **(6.76)**	**7,509** **(7.67)**	**7,917** **(8.09)**	**8,275** **(8.45)**	**10,720** **(10.95)**	**13,745** **(14.04)**	**20,029** **(20.46)**	**16,827** **(17.19)**	**97,896** **(100)**

Table 14. Number of Patients by AJCC Stage and Histologic Stage.

	Histologic Stage				
AJCC Stage	Localized	Regional	Distant	Unstaged	Total (%)
0	3950	.	.	.	3,950 (4.03)
I	9945	.	.	.	13,759(14.05) Stage I
IA	3814	.	.	.	
II	7750	.	.	.	15,813(6.15) Stage II
IIA	3120	.	.	.	
II-UNDEFINED SUFFIX	4943	.	.	.	
III	.	9936	.	.	11,669(11.92) Stage III
IIIA	.	383	.	.	
IIIC	.	1350	.	.	
IV	.	2974	7676	221	15,458(15.79) Stage IV
IVA	.	471	.	3	
IVB	.	474	.	.	
IVC	.	.	3639	.	
UNSTAGED	23803	.	.	13055	37,247(38.05) Unstaged
Error Condition	.	.	.	389	
TOTAL (%)	57,325 (58.56)	15,588 (15.92)	11,315 (11.56)	13,668 (13.96)	97,896 (100)

Table 15. AJCC Unstaged Patients By Histologic Stage and Year of Diagnosis.

	Year of Diagnosis									
Histologic Stage	1985	1986	1987	1988	1989	1990	1991	1992	1993	Total (%)
LOCALIZED	.	.	.	2,601	2,509	3,065	3,776	6,548	5,304	23,803 (63.91)
UNSTAGED	646	741	799	949	967	1,615	2,279	2,991	2,457	13,444 (36.09)
Total (%)	646 (1.73)	741 (1.99)	799 (2.51)	3,550 (9.53)	3,476 (9.33)	4,680 (12.56)	6,055 (16.26)	9,539 (25.6)	7,761 (20.84)	37,247 (100)

3.0 Description of Subjects

3.1 Prostate Cancer Case Selection

Figure 1 displays the case selection algorithm. Beginning with all prostate cancer cases included in the PEDSF (n=126,756), application of exclusion criteria eventually resulted in 97,896 patients being selected. Tables 16 and 17 display the diagnosis context and case number for the selected prostate cancer cases. Among the total 126,756 patients from the PEDSF, only one patient had 6 cases/diagnoses, which is the highest number of cases/diagnoses occurring in the file. Of the 126,756, only 84 had prostate cancer listed as the 4 th or later diagnosis/case. Two of those 84 patients were undetermined because of their unstaged cancers in the 1st and 2nd diagnoses. All others were excluded because their prostate cancers were not the first primary or one of the first primary malignant cancers. The highest number of diagnoses among the selected 97,896 patients is 5 and only two patients have 5 diagnoses (Table 16). As shown in Table 17, 98.60% (96,522 patients) of these selected patients were selected because their first cancer diagnosis/case was prostate cancer. Of those 96,522 patients selected because prostate cancer was their first cancer diagnosis/case, only 854 (=829+25) had one or more simultaneous diagnoses of other cancer. With respect to type of selection (diagnosis context), 97.7% (95,668 patients) of the selected patients had prostate cancer as their (only) first' primary cancer diagnosis, 1.80% had simultaneous cancer diagnoses, and 0.46% (449 patients) had some in situ cancer diagnosed before their prostate cancer (Table 17).

Figure 1. Prostate Cancer Case Selection.

Inclusion/Exclusion	Number of patients
PEDSF	Starting 126,756 patients
The first two diagnoses month are unknown	Excluded 6
Age at diagnosis<65,or Diagosed before 1985	Excluded 19,138
1st Sequence number not ('00', '01')	Excluded 2,368
1st primary cancer site is not Prostate	Excluded 7,348
Remaining patients	Included 97,896

Figure Footnotes:

(1) **Insitu vs. Malignant Cancer**: A tumor is in situ if the histologic stage has value '0', and otherwise is it is invasive.

(2) **Type of selections (diagnosis context)**: All the selected patients had prostate cancer as the first malignant or one of the first malignant cancers. Therefore there are five types of selections: *(i) In situ Prostate Before Malignant Prostate (ii) Other In situ Before Malignant Prostate (iii) Simultaneous Other Malignant with Malignant Prostate (iv) Simultaneous In situ with Malignant prostate (v) A combination of the above.*

(3) **Sequence Number**: Examination of sequence numbers revealed that not all first listed cancers in the records were the first diagnosed. Missing first diagnosed cancers were non-prostate cancer diagnoses and were cleaned out by the State Registry before sending the data to the National SEER. All the selected prostate cancer diagnoses are (i) with sequence number '00' or '01', or (ii) simultaneous with some diagnosis of such sequence numbers, or (iii) subsequent to some In situ cancer of such sequence number.

(4) **Unknown Month of Diagnosis**: When determining whether diagnoses are simultaneous, unknown month of diagnosis presents a problem. Fortunately, all relevant records (with at least one prostate cancer diagnosed) did have valid years of diagnosis. There were 550 records with unknown month of diagnosis and 26 of them were problematic due to multiple diagnoses in a same year. Whether these 26 records constitute "first primary" diagnoses cannot be determined and were excluded.

Table 16. Number of Patients By Number of SEER Cancer Diagnoses and the Case Number of Their Prostate Cancer.

Number of SEER Cancer Diagnoses	Listed Order (Case Number) of the Selected Prostate Cancer Diagnosis				Total (%)
	1st case	2nd case	3rd case	4th case	
01	89334	.	.	.	89,335 (91.25)
02	6751	1245	.	.	7,995 (8.17)
03	416	96	25	.	537 (0.55)
04	19	5	3	1	28 (0.03)
05	2	.	.	.	2 (0.00)
Total (%)	96,522 (98.60)	1,346 (1.3&)	28 (0.03)	1 (0.00)	97,896 (100)

Table 17. Number of Patients By Type of Selection and the Listed Order of the Selected Prostate Cancer Diagnosis.

	Listed order (case number) of the selected diagnosis				
	1st case	**2nd Case**	**3rd Case**	**4th case**	**Total(%)**
1st Primary	95,668	.	.		**95,668 (97.73)**
Insitu Prostate Before Malignant Prostate	.	2	.		**2 (0.00)**
Other Insitu Before Malignant Prostate	.	444	4	1	**449 (0.46)**
Simultaneous Malignant with Malignant Prostate	829	890	18		**1,737 (1.77)**
Simultaneous Insitu with Malignant prostate	25	6	.		**31 (0.03)**
Combination of the above	.	3	6		**9 (0.01)**
Total(%)	**96,522 (98.60)**	**1345 (1.37)**	**28 (0.03)**	**1 (0.00)**	**97,896 (100)**

3.2 Demographic Information

As shown in Table 18, 84.37% of the 97,894 patients are white, 9.21% are black, and the remaining 6.59% are other minority races. From Table 19, 91.57% of the selected patients were diagnosed with prostate cancer before they were 85 years old and 54.59% were diagnosed from age 65 to age 75. The number of cases selected from each SEER registry by race is displayed in Table 20. Detroit was the registry with the largest number of cases selected. Race/ethnic groups are not equally distributed across the registries, with Iowa, Utah, and Washington having a smaller percent of race/ethnic minority groups, and Hawaii the highest percentage of Asian Americans. Table 21 displays the number of prostate cancer cases selected within histologic stage by SEER registry.

Table 18. Number of Cases Selected By Age Category and Race.

Age	White	Black	Native American	Asian	Other	Total
65<<69	21020	2729	118	708	574	25,149 (25.69)
70<<74	24127	2600	98	881	685	28,391 (29.00)
75<<79	19274	1938	76	865	632	22,785 (23.27)
80<<84	11246	1043	46	604	385	13,324 (13.61)
85<<89	5086	491	25	304	146	6,052 (6.18)
90<<94	1504	174	15	77	27	1,797 (1.84)
95<<99	307	33	1	11	4	356 (0.36)
100<<	32	6	2	2	.	42 (0.04)
Total	82,596 (84.37)	9,014 (9.21)	381 (0.39)	3,452 (3.53)	2,453 (2.51)	97,896 (100)

Table 19. Number of Cases Selected By Age Category and Year of Diagnosis.

	Year of Diagnosis									
Age	**1985**	**1986**	**1987**	**1988**	**1989**	**1990**	**1991**	**1992**	**1993**	**Total (%)**
65<<69	1476	1592	1772	1871	2094	2665	3402	5379	4898	25,149 (25.69)
70<<74	1657	1752	2060	2220	2315	2972	4076	6090	5249	28,391 (29.00)
75<<79	1443	1551	1845	1868	1973	2522	3303	4633	3647	22,785 (23.27)
80<<84	965	1040	1088	1211	1165	1535	1839	2552	1929	13,324 (13.61)
85<<89	497	481	526	536	533	750	860	1049	820	6,052 (6.18)
90<<94	169	161	185	185	153	226	222	263	233	1,797 (1.84)
95<<99	51	33	32	21	38	46	36	55	44	356 (0.36)
100<<	2	4	1	5	4	4	7	8	7	42 (0.04)
Total (%)	6,260 (6.39)	6,614 (6.76)	7,509 (7.67)	7,917 (8.09)	8,275 (8.45)	10,720 (10.95)	13,745 (14.04)	20,029 (20.46)	16,827 (17.19)	97,896 (100)

Table 20. Number of Cases Selected by Registry and Race.

		Race					
		White	**Black**	**Native**	**Asian**	**Other**	**Total (%)**
Registry	**San Francisco**	9906	1542	16	734	254	**12,452 (12.72)**
	Connecticut	10523	613	2	6	81	**11,225 (11.47)**
	Detroit	12330	3937	8	26	584	**16,885 (17.25)**
	Hawaii	1098	25	180	2009	195	**3,507 (3.58)**
	Iowa	12294	123	6	3	104	**12,530 (12.80)**
	New Mexico	5218	77	116	3	4	**5,417 (5.53)**
	Seattle	13636	323	44	181	500	**14,684 (15.00)**
	Utah	5799	20	4	22	77	**5,922 (6.05)**
	Atlanta	3968	1318	.	9	131	**5,426 (5.54)**
	San Jose	1413	41	.	91	67	**1,612 (1.65)**
	Los Angeles	6412	995	5	368	456	**8,236 (8.41)**
	Total (%)	**82,596 (84.37)**	**9,014 (9.21)**	**381 (0.39)**	**3,452 (3.53)**	**2,453 (2.51)**	**97,896 (100)**

Table 21. Number of Cases Selected By Registry and Histologic Stage.

		Histologic Stage				
		Localized	Regional	Distant	Unstaged	**Total (%)**
Registry	San Francisco	6699	2340	1755	1658	**12,452 (12.72)**
	Connecticut	7119	1187	1497	1422	**11,225 (11.47)**
	Detroit	10676	1827	2019	2363	**16,885 (17.25)**
	Hawaii	2254	553	512	188	**3,507 (3.58)**
	Iowa	7143	2235	1787	1365	**2,530 (2.80)**
	New Mexico	3890	775	524	229	**5,417 (5.53)**
	Seattle	8185	2863	1246	2390	**14,684 (15.00)**
	Utah	3011	1164	599	1148	**5,922 (6.05)**
	Atlanta	2887	747	659	1133	**5,426 (5.54)**
	San Jose	849	320	134	309	**1,612 (1.65)**
	Los Angeles	4613	1577	583	1463	**8,236 (8.41)**
	Total (%)	**57,325 (58.36)**	**15,588 (15.92)**	**11,315 (11.56)**	**13,668 (13.96)**	**97,896 (100)**

3.3 Neoplasm Description

Table 22 displays the distribution of histologic stage by diagnosis year. The table demonstrates that the percent diagnosed at localized and regional stages has increased slightly over time, the percent diagnosed at distant stage has decreased and the percent unstaged has increased. This reflects the earlier detection of prostate cancers in the PSA testing era. Histologic stage by grade is displayed in Table 23. Among the 57,325 histologically localized staged patients, 30.54% (17,506 patients) had well differentiated tumors, 47.94% (27,494 patients) had moderately differentiated tumors, 14.77% (8,469 patients) had poorly differentiated tumors, and the remaining 6.75% had either un-differentiated or grade unknown tumors. In contrast, regional and distant tumors were much less likely to be well-differentiated and distant and unknown stage tumors were much more likely to have unknown grade.

Table 22. Number of Cases Selected By Histologic Stage and Year of Diagnosis.

Histologic Stage	Year of Diagnosis (column %)									
	1985	1986	1987	1988	1989	1990	1991	1992	1993	Total
Localized	3594 (57.41)	3773 (57.05)	4510 (60.06)	4705 (59.43)	4778 (57.75)	6082 (56.74)	7744 (56.34)	11931 (59.57)	10208 (60.66)	57,325 (58.56)
Regional	811 (12.96)	874 (13.21)	993 (13.22)	1039 (13.12)	1284 (15.52)	1670 (15.58)	2373 (17.26)	3561 (17.78)	2983 (17.73)	15,588 (15.92)
Distant	1209 (19.31)	1225 (18.52)	1205 (16.05))	1202 (15.18)	1220 (14.74)	1313 (12.25)	1309 (9.52)	1487 (7.42)	1145 (6.80)	11,315 (11.56)
Unstaged	646 (10.32)	742 (11.22)	801 (10.67)	971 (12.26)	992 (11.99)	1655 (15.44)	2319 (16.87)	3051 (15.23)	2491 (14.80)	13,668 (13.96)
Total (row %)	**6,260 (6.39)**	**6,614 (6.76)**	**7,509 (7.67)**	**7,917 (8.09)**	**8,275 (8.45)**	**10,720 (10.95)**	**13,745 (14.04)**	**20,029 (20.46)**	**16,827 (17.19)**	**97,896 (100)**

Table 23. Number of Cases Selected By Grade and Histologic Stage.

		HISTOLOGIC STAGE (Column Percent)				
		Localized	**Regional**	**Distant**	**Unstaged**	**Total (%)**
GRADE	**I: Well Differentiated**	17506 (30.54)	1196 (7.67)	667 (5.89)	1930 (14.12)	**21,2 (21.7**
	II: Moderately Differentiated	27494 (47.96)	8354 (53.59)	3266 (28.86)	5292 (38.72)	**44,4 (45.3**
	III: Poorly Differentiated	8469 (14.77)	5113 (32.80)	4694 (41.48)	2423 (17.72)	**20,6 (21.1**
	IV:Un-differentiated	374 (0.65)	240 (1.54)	323 (2.85)	123 (0.90)	**1,0 (1.0**
	Grade unknown	3482 (6.07)	685 (4.39)	2365 (20.93)	3900 (28.53)	**10,4 (10.6**
	TOTAL (%)	**57,325** (58.56)	**15,588** (15.92)	**11,315** (11.56)	**13,668** (13.96)	**97,8 (10**

4.0 Treatment Choice

4.1 Definition of Aggressive and Conservative Treatment

We classified patients into two main groups according to the SEER first course of treatment data: aggressive and conservative. Aggressive treatment was further categorized as radical prostatectomy or radiation treatment. There were 5,881 patients with unknown SEER first course of treatments. Of these we determined that 453 were treated aggressively, because procedure codes for radical prostatectomy or radiation treatment were found in the Medicare MEDPAR claims within four months of the diagnosis date (the period SEER defines as the first course of treatment). Of the prostate cancer cases selected, 5,428 (5.54%; 5,881 - 453) had unknown treatment type, 46,043 (47.03%) were conservatively treated and 46,425 (47.42%) were aggressively treated (Table 24.)

Table 25 displays the particular SEER surgery codes and radiation treatment codes used to define conservative and aggressive treatment categories. The aggressive treatment group includes 46,425 patients (where 46,425=45,972+453, and 45,972=18,528+28,858-1,414 from Table 25), the unknown group includes 5,428 patients (where 5,428=5,881-453 identified from MEDPAR and 5,881=1,996+7,359-743-163-2,568 from Table 25), and the conservative treatment group includes 46,043 patients (40,043=97,896-46,425-5,428).

Table 24. Number of Cases Selected By Treatment Type and Registry.

Registry	Treatment Type			Total (column %)
	Unknown	Conservative	Aggressive	
San Francisco	912 (7.32)	5769 (46.33)	5771 (46.35)	12,452(12.72)
Connecticut	685 (6.10)	5916 (52.70)	4624 (41.19)	11,225(11.47)
Detroit	1192 7.06)	7639 (45.24)	8054 (47.70)	16,885 (17.25)
Hawaii	180 (5.13)	1679 (47.88)	1648 (46.99)	3,507 (3.58)
Iowa	718 (5.73)	6802 (54.29)	5010 (39.98)	12,530 (12.80)
New Mexico	266 (4.91)	2585 (47.72)	2566 (47.37)	5,417 (5.53)
Seattle	687 (4.68)	5687 (38.73)	8310 (56.59)	14,684 (15.00)
Utah	225 (3.80)	2636 (44.51)	3061 (51.69)	5,922 (6.05)
Atlanta	281 (5.18)	2562 (47.22)	2583 (47.60)	5,426 (5.54)
San Jose	61 (3.78)	753 (46.71)	798 (49.50)	1,612 (1.65)
Los Angeles	221 (2.68)	4015 (48.75)	4000 (48.57)	8,236 (8.41)
Total (row%)	5,428(5.54)	46,043 (47.03)	46,425(47.42)	97,896(100)

Table 25. Number of Cases Selected By Whether Radical Prostatectomy Was Performed and Whether Radiation Treatment Was Received.

		Whether received Radiation Treatment			
		Not Received (Radiation Codes: 0/3/7)	Received (Radiation Codes: 1/2/4)	Unknown (Radiation Codes: 5/6/8/9)	Total (%)
Whether Surgery	No Surgery (Surgery Code: 00)	2,793	770	38	**3,601 (3.68)**
	Non surgical Diagnostic Method (Surgery codes: 01/02/03/04/05/06)	16,926	16,053	484	**33,463 (34.18)**
	Non Cancer-Directed rgery (Surgery Code: 07)	12	3	1	**16 (0.02)**
	Unknown if any Surgery done (Surgery Code: 09)	4,048	2,568	743	**7,359 (7.52)**
	TURP or Partial Prostatectomy (Surgery Codes: 10/20/30/38/40/80/90)	26,312	8,050	567	**34,929 (35.68)**
	Radical Prostatectomy (Surgery Codes: 50/58/60/68/70/78)	16,951	1,414	163	**18,528 (18.93)**
	Total	**67,042 (68.48)**	**28,858 (29.48)**	**1,996 (2.04)**	**97,896 (00)**

Footnotes for Table 25:

The site-specific surgical codes are:

'00'='00:No Surgical procedure'
'01'='01:Non-Cancer directed surgery'
'02'='02:Incisional/Needle/Asp_Biop of Primary'
'03'='03:Exploratory ONLY (No Biopsy)'
'04'='04:Bypass Surgery,-ostomy ONLY(No Biopsy)'
'05'='05:Expl ONLY & needle or aspiration biopsy of PRM/OTHE SITE'
'06'='06:Byp-Surg,-ostm ONLY & incis/ndle/asp biop of PRM/OTHER SITE'
'07'='07:Non-Cancer directed surgery, NOS'
'09'='09:Unknown if surgery done or diagnosed<=1982'
'10'='10:TURP,CRY-PRSTCTMY,LOC_SURG lesion-excis WITHOUT LYM_DISSEC'
'20'='20:TURP,CRY-PRSTCTMY,LOC_SURG lesion-excis WITH LYM_DISSEC'
'30'='30:Subtotal/simple PRSTCTMY WITHOUT LYM_DISSEC'
'38'='38:Subtotal/simple PRSTCTMY WITHOUT LYM_DISSEC(wth reconst-surg)'
'40'='40:Subtotal/simple PRSTCTMY WITH LYM_DISSEC'
'50'='50:Radical/total PRSTCTMY WITHOUT LYM_DISSEC'
'58'='58:Radical/total PRSTCTMY WITHOUT LYM_DISSEC(wth reconst-surg)'
'60'='60:Radical/total PRSTCTMY WITH LYM_DISSEC'
'68'='68:Radical/total PRSTCTMY WITH LYM_DISSEC(wth reconst-surg)'
'70'='70:CYS-PRSTCTMY,RAD_PRSTCTMY,PELV-EXENTER WTH/WTHOU LYM_DISSEC'
'78'='78:CYS-PRSTCTMY,RAD_PRSTCTMY,PELV-EXENTER WTH/WTHOU LYM_DISSEC'
'80'='80:Surg of regional and/or distant site(s)/node(s) ONLY'
'90'='90:PRSTCTMY,NOS;Surgery,NOS (including surg<=1982)'

The Radiation treatment codes are:

'0'='0:None'; '1'='1:Beam radiation'; '2'='2:Radioactive implants'; '3'='3:Radioisotopes'; '4'='4:BEAM & radio-implants/isotopes'; '5'='5:Radiation,NOS'; 6'='6:Undefined'; '7'='7:Pat/guardian refused rad-thrp'; '8'='8:Radiation-recomm,ukn if admin'; '9'='9:Unknown'

4.2. Factors Associated with Treatment Choice

Race and age were significantly associated with type of treatment chosen (p's less than 0.0001). Whites and Asians were more likely to receive aggressive treatment (Table 26) and the percent treated aggressively declined with advancing age. Among men aged 80 and over, less than 20 percent were treated aggressively. In contrast, two-thirds of men aged 65 to 69 were treated aggressively (Table 27).

Table 26. Number of Cases By Treatment Choice and Race.

Treatment Type	Race (row%)					
	White	Black	Native	Asian	Other	Total (column%)
Unknown	4456 (5.39)	664 (7.37)	33 (8.66)	173 (5.01)	102 (4.16)	**5,428 (5.54)**
Conservative	37537 (45.45)	5013 (55.61)	216 (56.69)	1710 (49.54)	1567 (63.88)	**46,043 (47.03)**
Aggressive	40603 (49.16)	3337 (37.02)	132 (34.65)	1569 (45.45)	784 (31.96)	**46,425 (47.42)**
Total (row%)	**82,596 (84.37)**	**9,014 (9.21)**	**381 (0.39)**	**3,452 (3.53)**	**2,453 (2.51)**	**97,896 (100)**

Table 27. Number of Cases by Treatment Choice and Age at Diagnosis.

Treatment Type	Age at Diagnosis (column %)								
	65<<69	70<<74	75<<79	80<<84	85<<89	90<<94	95<<99	>>100	Total (%)
Unknown	829 (3.30)	1118 (3.94)	1399 (6.14)	1064 (7.98)	664 (10.97)	270 (15.03)	77 (21.63)	7 (16.67)	**5,428 (5.54)**
Conservative	7590 (30.18)	10413 (36.68)	11810 (51.83)	9595 (72.01)	4888 (80.77)	1440 (80.13)	272 (76.40)	35 (83.33)	**46,043 (47.03)**
Aggressive	16730 (66.52)	16860 (59.39)	9575 (42.03)	2666 (20.01)	500 (8.26)	87 (4.84)	7 (1.97)	.	**46,425 (47.42)**
Total (row %)	**25,149 (25.69)**	**28,391 (29.00)**	**22,785 (23.27)**	**13,324 (13.61)**	**6,052 (6.18)**	**1,797 (1.84)**	**356 (0.36)**	**42 (0.04)**	**97,896 (100)**

Tumor grade and histologic stage were significantly associated with type of treatment chosen (p's less than 0.0001). Aggressively treated patients were more likely to have been diagnosed with a localized or regional tumor than were conservatively treated patients, and less likely than conservatively treated patients to have been diagnosed with a distant or unstaged tumor (Table 28). Aggressively treated patients were much more likely to have grade II (moderately differentiated) tumors than were conservatively treated patients. However, aggressively treated patients were less likely to have a grade I (well-differentiated) tumor (Table 29).

Table 28. Number of Cases by Treatment Choice and Histologic Stage.

Treatment Type	Histologic Stage				Total (column%)
	Localized	Regional	Distant	Unstaged	
Unknown	1686 (2.94)	396 (2.54)	1607 (14.20)	1739 (12.72)	5,428 (5.54)
Conservative	26501 (46.23)	3537 (22.69)	7840 (69.29)	8165 (59.74)	46,043 (47.03)
Aggressive	29138 (50.83)	11655 (74.77)	1868 (16.51)	3764 (27.54)	46,425 (47.42)
Total (row%)	57,326 (58.36)	15,588 (15.92)	11,315 (11.56)	13,668 (13.96)	97,896 (100)

Table 29. Number of Cases by Treatment Type and Tumor Grade.

Treatment Type	Tumor Grade (column percent)					
	I: Well differentiated	II: Moderately differentiated	III: Poorly differentiated	IV: Un-differentiated	Grade Unknown	Total (coumn %)
Unknown	805 (3.78)	1542 (3.47)	1131 (5.46)	85 (8.02)	1,865 (17.88)	5,428 (5.54)
Conservative	12690 (59.58)	16913 (38.09)	10034 (48.48)	588 (55.47)	5,818 (55.77)	46,043 (47.03)
Aggressive	7804 (36.64)	25951 (58.44)	9534 (46.06)	387 (36.51)	2,749 (26.35)	46,425 (47.42)
Total (row %)	21,299 (21.76)	44,406 (45.36)	20,699 (21.14)	1,060 (1.08)	10,432 (10.66)	97,896 (100)

Treatment type was also strongly associated with presence of co-morbidities. Of the total 341,665 records in the MEDPAR file, 3,796 had invalid discharge or admission dates. Therefore the information from these records was not used in the comorbidity analysis. The CCHPR coding scheme that maps ICD-9 codes into clinically meaningful groups was used for the comorbidity analysis. Thirteen conditions (13 CCHPR codes) were considered to be complications and therefore were considered to be a comorbidity only if they occurred in stays before the initial stay (the initial stay was the first stay following the prostate cancer diagnosis) or before the prostate cancer diagnosis date. Twenty-seven CCHPR codes represented other cancers and were excluded. Instead, we substituted an indicator from the SEER data for whether a patient had another malignant cancer simultaneous to the prostate cancer. Table 30 illustrates that patients who had another cancer diagnosed at the same time as the prostate cancer were much less likely to be treated aggressively ($p < 0.0001$). Patients who were treated aggressively had fewer co-morbid conditions detected from the MEDPAR data (Table 31). Age at diagnosis was also strongly associated with number of co-morbid conditions (Table 32).

Table 30. Number of Cases Selected By Treatment Type and Whether Other Cancers Were Diagnosed Simultaneously.

Treatment Type	Whether Prostate Cancer Was Diagnosed Simultaneously With Other Cancer		Total (column %)
	No	Yes	
Unknown	5,279 (5.49)	149 (8.58)	5,428 (5.54)
Conservative	44,934 (46.73)	1,109 (63.85)	46,043 (47.03)
Aggressive	45,946 (47.78)	479 (27.58)	46,425 (47.42)
Total (row %)	96,159 (98.23)	1,737 (1.77)	97,896 (100)

Table 31. Number of Cases Selected By Number of Co-morbid Conditions and Treatment Type.

Number of Conditions	Treatment Type			
	Unknown	Conservative	Aggressive	Total (%)
0==ILL	2468 (36.57)	16840 (45.47)	23465 (50.54)	42,773 (43.69)
1<=ILL<=2	955 (20.50)	9437 (17.59)	11066 (23.84)	21,458 (21.92)
3<=ILL<=4	1175 (25.48)	11730 (21.65)	8633 (18.60)	21,538 (22.00)
5<=ILL<=6	417 (8.81)	4056 (7.68)	2002 (4.31)	6,475 (6.61)
7<=ILL	413 (8.64)	3980 (7.61)	1259 (2.71)	5,652 (5.77)
Total (%)	5,428 (5.54)	46,043 (47.03)	46,425 (47.42)	97,896 (100)

Table 32. Number of Cases Selected By Total Number of Co-morbid Conditions and Age.

Number of Conditions	Age at Diagnosis								Total (%)
	65<<69	70<<74	75<<79	80<<84	85<<89	90<<94	95<<99	>>100	
0==ILL	11752 (46.73)	12841 (45.23)	10263 (45.04)	5379 (40.37)	1972 (32.58)	466 (25.93)	93 (26.12)	7 (16.67)	42,773 (43.69)
1<=ILL<=2	6609 (26.28)	6612 (23.29)	4516 (19.82)	2353 (17.66)	998 (16.49)	307 (17.08)	60 (16.85)	3 (7.14)	21,458 (21.92)
3<=ILL<=4	4825 (19.19)	6018 (21.20)	4989 (21.90)	3285 (24.65)	1726 (28.52)	564 (31.39)	116 (32.58)	15 (35.71)	21,538 (22.00)
5<=ILL<=6	1132 (4.50)	1607 (5.66)	1594 (7.00)	1182 (8.87)	679 (11.22)	230 (12.80)	43 (12.08)	8 (19.05)	6,475 (6.61)
7<=ILL	831 (3.30)	1313 (4.62)	1422 (6.24)	1126 (8.45)	677 (11.19)	230 (12.80)	44 (12.36)	9 (21.43)	5,652 (5.77)
Total (row %)	25,149 (25.69)	28,391 (29.00)	22,785 (23.27)	13,324 (13.61)	6,052 (6.18)	1,797 (1.84)	356 (0.36)	42 (0.04)	97,896 (100)

5.0 Instrumental Variables

5.1. Overview of Instrumental Variable Estimation Techniques

Instrumental variable (IV) estimation initially involves specifying a set of instrumental variables or "instruments" that are suitable for the research question at hand. In medical outcomes research, variables must satisfy the following two criteria to be suitable instruments : (1) the variable must be related to the possibility of patients receiving a particular treatment; and (2) the variable must have no effect on outcomes either directly or indirectly (e.g., through relationships with unmeasured confounding factors such as patient severity and unrecorded treatments). The first criterion is necessary to observe treatment variation across patients grouped by the instrument and can be established by analysis of the available data. The second criterion is necessary to insure that treatment variation observed from grouping patients using the instrument is not related to confounding factors such as patient severity. Because many confounders are unmeasured, the second criterion must remain an assumption. Consequently, researchers must build a strong theoretical case for acceptance of the validity of the second criterion. Estimated correlations between instruments and measured confounders may be used to bolster the case.

If a single instrument is used that divides patients into two groups, treatment effects can be estimated through a simple comparison of treatment and outcome rates across the two groups. IV analysis is more powerful, though, if several instruments are used and comparisons are made simultaneously across many patient groups defined by the instruments. Two-stage least squares (2SLS) has been shown to be the optimal method to combine the effects of several instruments in a single analysis. Each treatment decision in this study will be specified using the following two equation format and estimated using 2SLS:

Treatment Choice Equation: $T_i = \alpha + \gamma_1 * A_i + \gamma_2 * G_i + \gamma_3 * C_i + \gamma_4 * I_i + \varepsilon_i + \theta_i$

Outcome Equation: $O_i = \delta + \beta_1 * A_i + \beta_2 * G_i + \beta_3 * C_i + \beta_4 * T_i + \nu_i + \theta_i$

where:

O_i	=	1 if health outcome occurs (e.g. mortality within a time interval, re-treatment within a time interval), 0 otherwise. Cost equations will use total patient health care costs within the given time interval;
A_i	=	measured patient demographic characteristics;
G_i	=	measured tumor characteristics;
C_i	=	a set of binary variables based representing patient co-morbidities;
T_i	=	a binary variable equal to 1 if a patient received a specified treatment, 0 otherwise;
θ_i	=	unmeasured "confounding variables" that are related to both choice of treatment and outcomes;
ε_i, ν_i	=	the net impact of unmeasured variables that distinctly affect treatment choices and health outcome, respectively;
I_i	=	a set of binary variables that group patients according to values of instrumental variables that affect outcomes only through their impact on treatment choice.

Our treatment variable T_i is a binary variable indicating whether the patient was treated. The objective is to obtain unbiased estimates of β_4. Because "θ" is in both the treatment and outcome equations, the estimate of the treatment choice parameter in equation (2) will be biased if ordinary least squares (OLS) is applied. In the first stage of the estimation procedure, the treatment choice equation (i.e., equation (1)) is estimated using ordinary least squares. Equation 1 includes a set of binary variables, I_i, that group patients based on the value of each patient's instruments. The predicted values of treatment probabilities from the first stage regressions for each patient, "T-hat" are then substituted for T_i in equation (2). In the second stage, equation (2) is estimated using OLS. Because A_i and G_i and C_i are specified in both equations, the only source of variation in T-hat used to estimate β_4 is the variation in treatment rates across patient groups defined by the instruments. In addition, because we assumed that the instruments are unrelated to the unmeasured confounding factors "θ", the estimate of β_4 that results from this process will be unbiased and attributable only to treatment rate differences across patients grouped by the set of instruments.

5.2 Defining the Candidate Instrumental Variables

Following previous treatment outcome research using IV estimation, the candidate instruments under investigation in this research are based on differential patient access to various provider types. Separate instruments are being developed that are related to patient access to hospitals, radiation treatment facilities, and physicians that have different practice styles. The development of each of these instruments required the zip code of each patient from the SEER-Medicare Patient Entitlement and Diagnosis Summary File (PEDSF). In addition, the longitudes and latitudes of all zip code centroids were required to calculate distances. These data were obtained from Dr. Gerry Rushton in the Department of Geography at the University of Iowa. The processes used to estimate individual instrument values for each patient in our sample are described below.

From each SEER registry we obtained a list of all radiation treatment centers providing service in the region containing each registry. The lists contained the zip code of each center and years between 1984 and 1995 that each center provided services. A longitude and latitude was assigned to each radiation treatment center based on the centroid of the zip code containing the center. Using the longitude and latitudes of both the radiation treatment and each patient, the straight-line distance was calculated from each patient to the nearest radiation treatment center providing service in the year the patient was diagnosed. The median distance from the patients in our sample to their closest radiation treatment center was 6.6 miles. This distance was used as the cutoff value to group patients in our preliminary instrument validity comparisons.

Insufficient information existed to characterize the practice patterns of individual physicians using Medicare claims. Instead we opted to estimate the practice patterns of all physicians providing care to early-stage prostate cancer patients in the region surrounding each patient. To do so, we calculated the aggressive treatment rate for early-stage prostate cancer patients living within a twenty-mile radius around each patient's zip code in their diagnosis year. In preliminary instrument validity comparisons, patients were divided into two groups based whether they lived in a region with an aggressive treatment rate above or below fifty percent. In the future we plan to adjust these observed rates for differences in age, stage and grade and group patients based upon the ratio of actual to adjusted rates in their region.

Differential distance measures from patient residence to the nearest provider of a given classification has been a mainstay in instrument development in early IV health care outcomes research. For a given provider classification, differential distances can be computed for each patient using the following formula:

$(D_c - D_{nc})$ where

D_c = the distance from patient residence to the nearest provider of a given classification; and

D_{nc} = the distance from patient residence to the nearest provider not in the classification.

Patients with smaller differential distances live relatively closer to a provider in the specified classification. Note that this measure does not use the distance to the provider that patients actually utilized, but rather distances to providers that were available to the patients. In this research we classified all the short-term hospitals in each SEER region each year by whether they performed radical prostatectomies on Medicare patients during that year or not. Medicare inpatient claims from the MEDPAR files were used for this classification. The zip code of the hospitals in each SEER region were found from the Medicare provider files. Using these classifications, differential distance measures are now being calculated for each patient in our sample.

5.3 Evaluating the Validity of the Candidate Instrumental Variables

To be suitable instruments, variables must satisfy the following two criteria: (1) be related to the possibility of patients receiving a particular treatment and (2) have no effect on outcomes either directly or indirectly (e.g. through relationships with unmeasured confounding factors such as patient severity and unrecorded treatments). Because many confounders are unmeasured, the second criterion must

remain an assumption. However, by comparing rates of *measured* confounders between groups of patients defined by the candidate instruments, we can provide evidence in support of the assumption. Table 33 compares the demographic and tumor characteristics between: (1) conservatively treated and aggressively treated patients; (2) patients who live close to radiation treatment centers vs. those who live far; and (3) patients who reside in an area with a low rate of aggressive treatment vs. those living in an area with a high rate of aggressive treatment. Table 34 compares these same groups with respect to prevalence of co-morbid conditions. The total number of patients in these tables is 85,359 which excludes patients with unknown treatment choice or in-valid zip codes.

Tables 33 and 34 show that patients who are aggressively treated are younger, have a higher tumor grade, earlier disease stage, and a lower prevalence of most co-morbidities than patients who are treated conservatively. In contrast, although patients who live close to radiation treatment centers have a higher rate of aggressive treatment than patients who live far from radiation treatment centers, these patient groups are much more similar to each other than are the patients grouped based on treatment choice. Hence, distance to radiation treatment centers is related to treatment choice (52.22% of those who live close receive aggressive treatment whereas 49.31% of those who live far) but is unrelated to most measured confounders. Note that statistically significant differences are still found but these are not clinically significant, e.g., the percent of tumors that are grade I differs by 1% (22.60% vs. 21.65%) which is statistically ($P<0.0001$) but not clinically significant. This is due to the extremely large sample size and resulting high power.

Similarly, when patients are grouped by the aggressive treatment rates of their residential area, patients living in a high rate area are much more likely to have aggressive treatment. There is also a slight difference between these groups in prevalence of some co-morbidities (Table 34). However, the two patient groups formed by this variable are more similar than are the two groups formed by actual treatment choice. The differences in comorbidity prevalence between areas with high vx. low aggressive treatment rate (Table 34) may be attributable to the observed small differences between these groups in age and tumor characteristics. In future analyses we will group patients according to the ratio of actual to expected aggressive treatment rate in the area, where the expected rate is that based on the age and tumor characteristics in the area.

Table 35 demonstrates that the distance to radiation treatment centers is related to the choice of radiation treatment vs. radical prostatectomy among patients treated aggressively (n=43,333), but is not related to most measured confounders.

Table 33. Comparisons of Demographic and Tumor, Characteristics of Patients Grouped by Treatment Choice and by Candidate Instrumental Variables.

		Treatment Choice		Candidate Instrumental Variables			
				Distance to Radiation Center [a]		Area Aggressive Trtmt. Rate [b]	
		Conservative	Aggressive	Close	Far	Low	High
Number of Patients		42,026 (49.93)	43,333 (50.07)	42,624 (49.53)	42,735 (50.07)	44,362 (51.97)	40,997 (48.03)
Treatment Choice							
	%Aggressive Treatment	0	100	52.22	49.31***	42.18	60.06***
Age at Diagnosis							
	%65-69	16.45	36.16 ***	26.60	26.31	25.05	27.97***
	%70-74	22.58	36.25***	29.56	29.48	28.21	30.94***
	%75-79	25.70	20.52***	23.32	22.83	23.34	22.79
	%80-84	20.81	5.80***	12.84	13.54*	14.41	11.87***
	%85-89	10.67	1.07***	5.72	5.89	6.60	4.93***
	%90-94	3.12	0.18***	1.61	1.64	1.96	1.27***
	%95-99	0.59	0.01***	0.30	0.30	0.39	0.20***
	%>100	0.07	0.00***	0.05	0.02 *	0.04	0.03
Grade							
	%I	27.65	16.76***	22.60	21.65***	24.89	19.13***
	%II	36.50	56.07***	45.54	47.32***	42.31	50.90***
	%III	21.90	20.65***	20.96	21.58*	22.02	20.45***
	%IV	1.33	0.87***	1.01	1.18*	1.36	0.81***
	%Unknown	12.61	5.66***	9.89	8.27***	9.42	8.72***
Histologic Stages							
	%Localized	56.97	65.52***	60.62	59.56	60.71	58.79***
	%Regional	7.76	25.75***	16.22	17.56***	15.11	18.82***
	%Distant	17.06	3.78***	10.60	10.03***	12.67	7.77***
	%Unstaged	18.21	7.95***	13.15	12.85	11.51	14.62***
Medicare							
	%With full Medicare	95.35	94.18***	94.73	94.78	95.19	94.29***

***, * Significantly different across groups at .99, .95 confidence levels, respectively. [a] Close patients that live less than median-distance from the closest radiation treatment center. [b] High patients that live in regions with Aggressive treatment rate greater than 50%.

Table 34. Comparisons of Co-morbidity Characteristics of Patients Grouped by Treatment Choice and by Candidate Instrumental Variables.

		Treatment Choice		Candidate Instrumental Variables			
				Distance to Radiation Center [a]		Area Aggressive Trtmt. Rate [b]	
		Conservative	Aggressive	Close	Far	Low	High
Other Cancer							
	%With other cancer	236	0.99***	1.69	1.64	1.92	1.39***
CCHP Code	Diseases - % With specified condition						
49	Diabetes mellitus without complication	5.478	3.937***	4.849	4.542*	5.085	4.273***
55	Fluid and electrolyte disorders	7.579	4.373***	6.051	5.852	5.892	6.015
59	Deficiency and other anemia	4.704	2.954***	3.859	3.772	4.078	3.532***
60	Acute posthemorrhagic anemia	2.318	7.279***	4.375	5.295***	3.429	6.359***
98	Essential hypertension	11.391	12.115***	11.787	11.730	11.717	11.803
101	Coronary atherosclerosis and other he	10.843	8.712***	9.994	9.528*	10.139	9.352***
105	Conduction disorders	3.710	2.091***	2.876	2.899	3.372	2.364***
106	Cardiac dysrhythmias	8.811	5.435***	7.052	7.142	7.635	6.515***
108	Congestive heart failure, nonhyperten	6.018	1.592***	3.789	3.753	4.091	3.425***
122	Pneumonia (except that caused by tube	3.522	1.126***	2.172	2.438 ***	2.459	2.139***
127	Chronic obstructive pulmonary disease	8.949	5.155***	6.759	7.287***	7.678	6.315***
143	Abdominal hernia	3.298	2.970***	2.832	3.430***	3.429	2.810***
159	Urinary tract infections	6.962	2.305***	4.772	4.425*	5.498	3.625***
161	Other diseases of kidney and ureters	3.222	1.161***	2.175	2.176	2.624	1.690***
162	Other diseases of bladder and urethra	14.146	5.495***	9.288	10.219***	11.934	7.396***
163	Genitourinary symptoms and ill-define	18.027	5.804***	11.747	11.897	14.436	8.993***
164	Hyperplasia of prostate	19.497	6.741***	13.244	12.800	16.316	9.457***
165	Inflammatory conditions of male genit	3.084	1.558***	2.163	2.455***	2.622	1.971***
238	Complications of surgical procedures	6.222	9.328***	7.310	8.286***	7.157	8.493***

***, * Significantly different across groups at .99, .95 confidence levels, respectively. [a] Close patients that live less than median-distance from the closest radiation treatment center. [b] High patients that live in regions with Aggressive treatment rate greater than 50%.

Table 35. Comparisons of Demographic and Tumor Characteristics of Patients Grouped by Treatment Choice and by Distance to Radiation Treatment Centers, Among Those Treated Aggressively.

		Treatment Choice		Distance to Radiation Center *	
		No Radiation	Radiation Trt	Close	Far
Number of Patients		16,468 (38.00)	26,865 (62.00)	22,260 (51.37)	21,073 (48.63)
Treatment Choice					
	% Radiation Treatment	0	100	65.13	58.68***
Age at Diagnosis					
	%65-69	51.91	26.51***	35.29	37.08***
	%70-74	37.49	35.49***	36.22	36.28
	%75-79	9.43	27.33***	21.09	19.93***
	%80-84	1.01	8.73***	5.99	5.59
	%85-89	0.14	1.65***	1.22	0.92***
	%90-94	0.02	0.28***	0.16	0.19
	%95-99	0	0.02	0.02	0
Grade					
	%I	13.70	18.64***	17.24	16.26***
	%II	63.68	51.40***	55.10	57.09***
	%III	19.83	21.15***	20.44	20.87
	%IV	0.61	1.02***	0.84	0.89
	%Unknown	2.18	7.79***	6.37	4.90***
Histologic Stage					
	%Localized	55.73	66.69***	63.76	61.22***
	%Regional	38.89	17.69***	23.99	27.60***
	%Distant	0.50	5.79***	4.09	3.45***
	%Unstaged	4.88	9.83***	8.16	7.73
Medicare					
	%With full Medicare	92.28	95.35***	94.25	94.12

Table 36. Comparisons of Co-morbidity Characteristics of Patients Grouped by Treatment Choice and by Distance to Radiation Treatment Center, Among Patients Treated Aggressively.

		Treatment Choice		Distance to Radiation Center *	
		No Radiation	Radiation Trt	Close	Far
Other Cancer					
	%With other cancer	0.89	1.05	1.04	0.93
CCHP Code	**Diseases - % wit specified condition**				
49	Diabetes mellitus without complicatio	4.682	3.480***	4.106	3.758
55	Fluid and electrolyte disorders	5.933	3.417***	4.245	4.508
59	Deficiency and other anemia	4.469	2.025***	3.010	2.895
60	Acute posthemorrhagic anemia	16.365	1.709***	6.312	8.300***
98	Essential hypertension	16.729	9.287***	12.107	12.125
101	Coronary atherosclerosis and other he	9.084	8.483*	8.931	8.480
105	Conduction disorders	2.144	2.058	2.129	2.050
106	Cardiac dysrhythmias	6.115	5.018***	5.494	5.372
127	Chronic obstructive pulmonary disease	5.453	4.973*	5.067	5.248
130	Pleurisy, pneumothorax, pulmonary col	2.909	0.998***	1.631	1.822
143	Abdominal hernia	4.627	1.954***	2.695	3.260***
145	Intestinal obstruction without hernia	2.131	0.815***	1.271	1.362
159	Urinary tract infections	1.882	2.565***	2.457	2.145*
162	Other diseases of bladder and urethra	4.214	6.280***	5.409	5.585
163	Genitourinary symptoms and ill-define	3.838	7.009***	6.083	5.509*
164	Hyperplasia of prostate	5.690	7.385***	6.882	6.591

238	Complications of surgical procedures	16.954	4.653***	8.720	9.970***
260	E (external causes of injury and pois	2.022	1.061***	1.366	1.490

Key Research Accomplishments

Key accomplishments this project year include:

- Acquiring, downloading, reading, and documenting the numerous files from the SEER-Medicare linked data for all prostate cancers from eleven SEER registries;
- Obtaining the zip code and years of operation for all radiation treatment centers in the eleven SEER areas;
- Locating and obtaining a detailed data dictionary that was not provided with the data. Researching the voluminous data dictionary to understand the Medicare files;
- Constructing and validating a case selection algorithm to apply the study inclusion and exclusion criteria;
- Evaluating data quality of key variables;
- Constructing treatment variables from the SEER data;
- Constructing candidate instrumental variables from the Medicare data and from locations of radiation treatment centers; and
- Providing preliminary evidence that patients grouped by access to care variables do have variations in treatment choice that is not likely explained by demographic, tumor, or comorbidity characteristics of the groups.

Reportable Outcomes

The SEER-Medicare linked database is very large and complex. During the first project year we have developed experience working with the data and have developed a library of programs and files. This will increase the efficiency of analyses in the remaining project period and for future projects.

We anticipate completing a manuscript on factors related to treatment choice for prostate cancer in the next three months.

Conclusions

In describing the factors that are related to the sorting of patients into conservative or aggressive treatments, we have found that there are variables that are related to whether a patient receives conservative vs. aggressive treatment for prostate cancer, but *not appreciably related to measured co-morbidity.* Thus we have provided evidence that the candidate instrumental variables that we proposed do satisfy the two criteria that an instrument must fulfill: (1) the variable must be related to the possibility of patients receiving a particular treatment; and (2) the variable must have no effect on outcomes either directly or indirectly (e.g., through relationships with unmeasured confounding factors such as patient severity and unrecorded treatments). If the candidate instruments had failed this test, it would not have been possible to proceed to Task 2, estimating unbiased treatment effects for conservative vs. aggressive treatment. In addition to the two candidate instruments that were examined, another candidate has been developed (differential distance to prostatectomy hospitals) and we will complete testing on this variable in the next few weeks.

The co-morbidity data were obtained from Medicare data; the SEER registries do not collect co-morbidity. This demonstrates the usefulness of the SEER-Medicare linked database for health outcomes research.

References

1. Wingo PA, Tong T, Bolden S. Cancer Statistics, 1995. CA Cancer J Clin 1995;45:8-30
2. Potosky AL, Miller BA, Albertsen PC, et al. The role of increasing detection in the rising incidence of prostate cancer. JAMA 1995;273:548-52.
3. Lu-Yao GL, Greenberg ER. Changes in prostate cancer incidence and treatment in USA. LANCET 1994;343:251-54.
4. Montie JE. Current prognostic factors for prostate carcinoma. Cancer, 1996;78:341-4.
5. Middleton RG, Thompson IM, Austenfeld MS, et al Prostate cancer clinical guidelines panel summary report on the management of clinically localized prostate cancer. J Urology, 1995;154:2144-8.
6. Adolfsson J. Radical prostatectomy, radiotherapy or deferred treatment for localized prostate cancer? Cancer Surveys 1995;23:141-8
7. Lu-Yao GL, Yao SL. Population-based study of long-term survival in patients with clinically localised prostate cancer. Lancet 1997;349:906-10.
8. Lu-Yao GL, Potosky AL, Albertsen PC et al. Follow-up prostate cancer treatments after radical prostatectomy: a population-based study. J Natl Cancer Inst 1996;88:166-73.
9. McClellan M, Newhouse JP. The marginal benefits of medical treatment intensity, Cambridge, Mass: National Bureau of Economic Research: Working Paper, 1993.
10. Wald A. The fitting of straight lines if both variables are subject to errors. Annals of Mathematical Statistics, 1940;11:284-300.
11. Durbin J. Errors in variables. Review of the International Statistics Institute. 1954;1:23-32.

12. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. National Bureau of Economic Research, Working Paper, 1993.

13. McClellan M, McNeil BJ, and Newhouse JP: Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality: analysis using instrumental variables. JAMA 1994;272(11):859-866.

14. Chow G. Test of Equality Between Sets of Coefficients in Two Linear Regressions. Econometrica. 1960;28:591-605.

15. Byar DP. Problems with using observational databases to compare treatments. Statistics in Medicine 1991:10:663.

16. Temple R. Problems in the use of large data sets to assess effectiveness. International Journal of Technology Assessment 1990;6:211.

17. Brooks JM, McClellan M, and Wong H: Evaluation of the marginal benefits of invasive treatment for acute myocardial infarction in a general population using instrumental variables techniques, ongoing research.

18. Brooks JM, Steiner C, and Wong H: Changes in the timing of bypass surgery following acute myocardial infarction, ongoing research.

19. Potosky AL, Riley GF, Lubitz JD, Mentnech RM, Kessler LG. Potential for cancer-related health services research using a linked Medicare-tumor registry database. Medical Care. 1993;31:732-48.

20. Friedman, M, A, Theory of the Consumption Function, 1957, Princeton, NJ, Princeton University Press.

21. Zellner, A. "Estimation of regression relationships containing unobserved independent variables", International Economic Review, 11, 1970, 441-454.

22. Angrist, J.D. "Grouped-data estimation and testing in simple labor-supply models", Journal of Econometrics, 47, 1991, 243-266.

23. Moffitt, R.A. "Selection bias adjustment in treatment effect models as a method of aggregation", National Bureau of Economic Research: 1995, Working Paper.

24. Amemyia T: Advanced Econometrics. Cambridge, MA, Harvard University Press, 1985.

25. SEER Program Code Manual. Bethesda (MD): National Cancer Institute, 1992.

DEPARTMENT OF THE ARMY
US ARMY MEDICAL RESEARCH AND MATERIEL COMMAND
504 SCOTT STREET
FORT DETRICK, MARYLAND 21702-5012

REPLY TO
ATTENTION OF:

MCMR-RMI-S (70-1y) 1 Apr 03

MEMORANDUM FOR Administrator, Defense Technical Information Center (DTIC-OCA), 8725 John J. Kingman Road, Fort Belvoir, VA 22060-6218

SUBJECT: Request Change in Distribution Statement

1. The U.S. Army Medical Research and Materiel Command has reexamined the need for the limitation assigned to technical reports written for this Command. Request the limited distribution statement for the enclosed accession document numbers be changed to "Approved for public release; distribution unlimited." Copies of these reports should be released to the National Technical Information Service.

2. Point of contact for this request is Ms. Judy Pawlus at DSN 343-7322 or by e-mail at judy.pawlus@det.amedd.army.mil.

FOR THE COMMANDER:

Encl

PHYLLIS M. RINEHART
Deputy Chief of Staff for
Information Management

ADB277986
ADB263450
ADB267669
ADB277564
ADB261754
ADB257280
ADB283722
ADB249627
ADB282841
ADB266235
ADB283529
ADB283519
ADB256683
ADB262564
ADB271045
ADB283537
ADB257204
ADB283513
ADB281571
ADB262777
ADB270818
ADB283748
ADB274588
ADB283788
ADB259015
ADB266031